PCT/GB 2004 / 0 0 3 7 7 3

INVESTOR IN PEOPLE

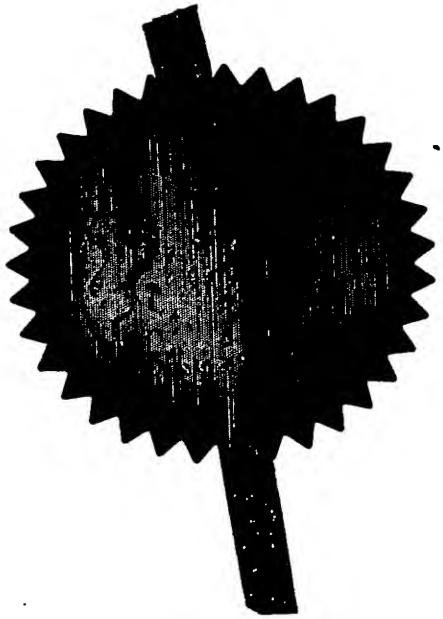
The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed Stephen Hordley

Dated 13 September 2004

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Patents Form 1/77

Patents Act 1977
(Rule 16)

THE PATENT OFFICE
-3 SEP 2003
RECEIVED BY HAND

The Patent Office

04SEP03 E834787-1 002890
P01/7700 0.00-0320648.9

1/77

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
South Wales
NP9 1RH

1. Your reference

JWJ01047GB

2. Patent application number
(The Patent Office will fill in this part)

03 SEP 2003 0320648.9

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

RANDOX LABORATORIES LTD.
Ardmore
Diamond Road
Crumlin
Co. Antrim
BT29 4QY

Patents ADP number *(if you know it)*

07420268001

If the applicant is a corporate body, give the country/state of its incorporation

UNITED KINGDOM

4. Title of the invention

MOLECULAR MARKER

5. Name of your agent *(if you have one)*

Gill Jennings & Every

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

Broadgate House
7 Eldon Street
London
EC2M 7LH

Patents ADP number *(if you know it)*

745002

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

YES

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	13
Claim(s)	1
Abstract	-
Drawing(s)	12

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	NO

11. For the applicant
Gill Jennings & Every

I/We request the grant of a patent on the basis of this application.

Signature John Jappy Date 3 September 2003

12. Name and daytime telephone number of person to contact in the United Kingdom

JAPPY, John William Graham
020 7377 1377

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

MOLECULAR MARKER

Field of the Invention

This invention relates to the detection of the presence of or the risk of cancer, in particular breast cancer.

Background of the Invention

There are over 1 million cases of breast cancer per year on a global basis, of which around 0.5 million are in the US, 40,000 are in the UK and nearly 2,000 in Ireland. It is the leading cause of cancer deaths among women. Although the overall incidence of the disease is increasing within the western world, wider screening and improved treatments have led to a gradual decline in the fatality rate of about 1% per year since 1991. Patients diagnosed with early breast cancer have greater than a 90% 5 year relative survival rate, as compared to 20% for patients diagnosed with distally metastasised breast cancer. Nonetheless, there is no definitive early-stage screening test for breast cancer, diagnosis currently being made on the results of mammography and fine needle biopsy. Mammography has its limitations, with over 80% of suspicious results being false positives and 10-15% of women with breast cancer providing false negative results. Often the tumour has reached a late stage in development before detection, reducing the chances of survival for the patient and increasing the cost of treatment and management for the healthcare system. More sensitive methods are required to detect small (<2 cm diameter) early stage *in-situ* carcinomas of the breast, to reduce patient mortality. In addition to early detection, there remain serious problems in classifying the disease as malignant or benign, in the staging of known cancers and in differentiating between tumour types. Finally, there is a need to monitor ongoing treatment effects and to identify patients becoming resistant to particular therapies. Such detection processes are further complicated, as the mammary gland is one of the few organs that undergo striking morphological and functional changes during adult life, particularly during pregnancy, lactation and involution, potentially leading to changes in the molecular signature of the same mammary gland over time.

Diagnosis of disease is often made by the careful examination of the relative levels of a small number of biological markers. Despite recent advances, the contribution of the current biomarkers to patient care and clinical outcome is limited. This is due to the low diagnostic sensitivity and disease specificity of the existing markers. Some molecular biomarkers, however, are being used routinely in disease diagnosis, for example prostate specific antigen in prostate cancer screening, and new candidate markers are being discovered at an increasing rate (Pritzker, 2002). It is becoming accepted that the use of a panel of well-validated biomarkers would enhance the positive predictive value of a test and minimize false positives or false negatives (Srinivas et al., 2002). In addition, there is now growing interest in neural networks, which show the promise of combining weak but independent information from various biomarkers to produce a prognostic/predictive index that is more informative than each biomarker alone (Yousef et al., 2002).

As more molecular information is collated, diseases such as breast cancer are being sub-divided according to genetic signatures linked to patient outcome, providing valuable information for the clinician. Emerging novel technologies in molecular medicine have already demonstrated their power in discriminating between disease sub-types that are not recognisable by traditional pathological criteria (Sorlie et al., 2001) and in identifying specific genetic events involved in cancer progression (Srinivas et al., 2002). Further issues need to be addressed in parallel, relating to the efficacy of biomarkers between genders and races, thus large scale screening of a diverse population is a necessity.

The management of breast cancer could be improved by the use of new markers normally expressed only in the breast but found elsewhere in the body, as a result of the disease. Predictors of the activity of the disease would also have valuable utility in the management of the disease, especially those that predict if a ductal carcinoma *in situ* will develop into invasive ductal carcinoma.

Summary of the Invention

According to a first aspect of the present invention, there is a method for the detection of the presence of or the risk of cancer in a patient, comprising the steps of:

(i) isolating a sample of the patient's genome; and

(ii) detecting the presence or expression of the gene characterised by the nucleotide sequence identified as SEQ ID No. 1, wherein the presence or expression of the gene indicates the presence of or the risk of cancer.

According to a second aspect of the invention, an isolated polynucleotide comprises the nucleotide sequence identified herein as SEQ ID No. 1, or its complement, or a polynucleotide of at least 15 consecutive nucleotides that hybridises to the sequence (or its complement) under stringent hybridising conditions.

According to a third aspect of the present invention, an isolated peptide comprises the sequence identified herein as SEQ ID No.2, or a fragment thereof of at least 10 consecutive amino acid residues.

According to a fourth aspect of the invention, an antibody has an affinity of at least 10^{-6}M for a peptide as defined above.

According to a fifth aspect of the invention, a polynucleotide that hybridises to or otherwise inhibits the expression of an endogenous DD11 gene, is used in the manufacture of a medicament for the treatment of cancer, in particular breast cancer.

Description of the Invention

The present invention is based on the identification of a gene that is expressed in a patient suffering cancer, in particular breast, uterus or testicular cancer. Identification of the gene (or its expressed product) in a sample obtained from a patient indicates the presence of or the risk of cancer in the patient.

The invention further relates to reagents such as polypeptide sequences, useful for detecting, diagnosing, monitoring, prognosticating, preventing, imaging, treating or determining a pre-disposition to cancer.

The methods to carry out the diagnosis can involve the synthesis of cDNA from mRNA in a test sample, amplifying as appropriate portions of the cDNA corresponding to the gene or a fragment thereof and detecting the product as an indication of the presence of the disease in that tissue, or detecting translation products of the mRNAs comprising gene sequences as an indication of the

presence of the disease.

Useful reagents include polypeptides or fragment(s) thereof which may be useful in diagnostic methods such as RT-PCR, PCR or hybridisation assays of mRNA extracted from biopsied tissue, blood or other test samples; or proteins which are the translation products of such mRNAs; or antibodies directed against these proteins. These assays also include methods for detecting the gene products (proteins) in light of possible post-translational modifications that can occur in the body, including interactions with molecules such as co-factors, inhibitors, activators and other proteins in the formation of sub-unit complexes.

The gene associated with cancer, is characterised by the polynucleotide shown as SEQ ID No. 1. The expressed product of the gene is identified herein by SEQ ID No. 2. Identification of the gene or its expressed product may be carried out using techniques known for the detection or characterisation of polynucleotides or polypeptides. For example, isolated genetic material from a patient can be probed using short oligonucleotides that hybridise specifically to the target gene. The oligonucleotide probes may be detectably labelled, for example with a fluorophore, so that, upon hybridisation with the target gene, the probes can be detected. Alternatively, the gene, or parts thereof, may be amplified using the polymerase chain reaction, with the products being identified, again using labelled oligonucleotides.

Diagnostic assays incorporating this gene, or associated protein or antibodies will include, but are not limited to:

Polymerase chain reaction (PCR)
Reverse transcription PCR
Real-time PCR
In-Situ hybridisation
Southern dot blots
Immuno-histochemistry
Ribonuclease protection assay
cDNA array techniques
ELISA

Protein, antigen or antibody arrays on solid supports such as glass or ceramics, useful in binding studies.

Small interfering RNA functional assays.

All of the above techniques are well known to those in the art.

The present invention is also concerned with isolated polynucleotides that comprise the sequence identified as SEQ ID No. 1, or its complement, or fragments thereof that comprise at least 15 consecutive nucleotides, preferably 30 nucleotides, more preferably at least 50 nucleotides. Polynucleotides that hybridise to a polynucleotide as defined above, are also within the scope of the invention. Hybridisation will usually be carried out under stringent conditions. Stringent hybridising conditions are known to the skilled person, and are chosen to reduce the possibility of non-complementary hybridisation. Examples of suitable conditions are disclosed in Nucleic Acid Hybridisation. A Practical Approach (B.D. Hames and S.J. Higgins, editors IRL Press, 1985). More specifically, stringent hybridisation conditions include overnight incubation at 42°C in a solution comprising: 50% formamide, 5 x SSC (150 mM NaCL, 15 mM trisodium citrate), 50 mM sodium phosphate (ph7.6), 5 x Denhardt's solution, 10% dextran sulphate and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing in 0.1 x SSC at about 65°C.

The identification of the DD11 gene also permits therapies to be developed, with the gene being a target for therapeutic molecules. For example, there are now many known molecules which have been developed for gene therapy, to target and prevent the expression of a specific gene. One particular molecule is a small interfering RNA (siRNA), which suppresses the expression of a specific target protein by stimulating the degradation of the target mRNA. Other synthetic oligonucleotides are also known which can bind to a gene of interest (or its regulatory elements) to modify expression. Peptide nucleic acids (PNAs) in association with DNA (PNA-DNA chimeras) have also been shown to exhibit strong decoy activity, to alter the expression of the gene of interest.

The present invention also includes antibodies raised against a peptide of the invention. The antibodies will usually have an affinity for the peptide of at least 10^{-6}M, more preferably, 10^{-9}M and most preferably at least 10^{-11}M. The

antibody may be of any suitable type, including monoclonal or polyclonal. Assay kits for determining the presence of the peptide antigen in a test sample are also included. In one embodiment, the assay kit comprises a container with an antibody, which specifically binds to the antigen, wherein the antigen comprises at least one epitope encoded by the DD11 gene. These kits can further comprise containers with useful tools for collecting test samples, such as blood, saliva, urine and stool. Such tools include lancets and absorbent paper or cloth for collecting and stabilising blood, swabs for collecting and stabilising saliva, cups for collecting and stabilising urine and stool samples. The antibody can be attached to a solid phase, such as glass or a ceramic surface.

Detection of antibodies that specifically bind to the antigen in a test sample suspected of containing these antibodies may also be carried out. This detection method comprises contacting the test sample with a polypeptide which contains at least one epitope of the gene. Contacting is performed for a time and under conditions sufficient to allow antigen/antibody complexes to form. The method further entails detecting complexes, which contain the polypeptide. The polypeptide complex can be produced recombinantly or synthetically or be purified from natural sources.

In a separate embodiment of the invention, antibodies, or fragments thereof, against the antigen can be used for the detection of image localisation of the antigen in a patient for the purpose of detecting or diagnosing the disease or condition. Such antibodies can be monoclonal or polyclonal, or made by molecular biology techniques and can be labelled with a variety of detectable agents, including, but not limited to radioisotopes.

In a further embodiment, antibodies or fragments thereof, whether monoclonal or polyclonal or made by molecular biology techniques, can be used as therapeutics for the treatment of diseases characterised by the expression of the gene of the invention. The antibody may be used without derivitisation, or it may be derivitised with a cytotoxic agent such as radioisotope, enzyme, toxin, drug, pro-drug or the like.

The term "antibody" refers broadly to any immunologic binding agent such as IgG, IgM, IgA, IgD and IgE. Antibody is also used to refer to any antibody-like

molecule that has an antigen-binding region and includes, but is not limited to, antibody fragments such as single domain antibodies (DABS), Fv, scFv etc. The techniques for preparing and using various antibody-based constructs and fragments are well known in the art.

If desired, the cancer screening methods of the present invention may be readily combined with other methods in order to provide an even more reliable indication of diagnosis or prognosis, thus providing a multi-marker test.

The following example illustrates the invention with reference to the accompanying drawings.

Example

A number of differentially expressed gene fragments were isolated from cDNA populations derived from matched clinical samples of breast cancer patients, using non-isotopic differential display (DDRT-PCR). One of these fragments, DD11 was revealed to be significantly up-regulated in breast tumour tissue samples from a number of donors. The expression profile of this novel molecular marker, its full length and corresponding presumed protein sequence is detailed herein.

Materials and methods.

We identified differential gene expression between matched pairs of normal mammary and tumour tissue from the same donor. Tissue samples were obtained, with full ethical approval and informed patient consent, from Pathlore, Peterborough, UK. Following the surgical removal of a tumour, one sample of the tumour tissue was collected, as was a sample from the adjacent, co-excised normal tissue. Messenger RNA was extracted and cDNA subsequently synthesised, using Dynal dT_{18}-tagged Dynabeads and Superscript II reverse transcription protocols, respectively. Differential display reverse transcription PCR (DDRT-PCR) was employed to observe differences between the gene expression profiles of these matched samples, and individual gene transcripts showing up- or down-regulation were isolated and investigated further.

First described by Liang & Pardee (1992), differential display reverse transcription PCR (DDRT-PCR) uses mRNA from two or more biological samples as templates for representative cDNA synthesis by reverse transcription, with

one of 3 possible anchor primers. Each of the 3 sub-populations was PCR-amplified using its respective anchor primer coupled with one of 80 arbitrary 13-mer primers. This number of primer combinations has been estimated to facilitate the representation of 96% of expressed genes in an mRNA population (Sturtevant, 2000). This population sub-division results in the reduction of the estimated 12,000-15,000 mRNAs expressed in eukaryotic cells to 100-150 transcripts by the end of second strand cDNA synthesis for each primer set. This facilitates the parallel electrophoretic separation and accurate visualization of matched primer sets on a polyacrylamide gel, leading to the identification of gene fragments expressed in one tissue sample but not the other.

Excision and re-amplification of fragments of interest was followed by removal of false positives through reverse Southern dot blotting. This entailed the spotting of each re-amplified fragment onto duplicate nylon membranes (Hybond N+, Amersham Pharmacia Biotech) and hybridising these with either the tumour or normal tissue cDNA population of the donor from which the fragments were derived. Those fragments confirmed as differentially expressed were then direct-sequenced, i.e. without cloning, followed by web-based database interrogation to determine if each gene was novel. Fragments not matching known genes were regarded as potentially representing novel markers for the breast cancer from which they were derived. Further screening of each transcript was performed by either semi-quantitative RT-PCR or real-time PCR, using a suite of matched cDNA populations from a number of breast tumour donors. In all cases, β-actin was used as a constitutive reference gene, for calibrating the cDNA templates and as an internal positive control during PCR. Expression of each putative novel marker gene was performed through the use of gene-specific primer sets on the calibrated templates. Full-length transcripts of the novel gene fragments, including the open reading frame (that piece of the gene that encodes the protein) were then synthesized using a complex process known as 5' RACE (rapid amplification of cDNA ends), which incorporates gene-specific extension and amplification, verifiable by sequencing.

Determination of tissue specificity was assayed using the gene-specific primers from each novel marker against cDNA populations from non-breast

tissue, including brain, heart, lymphocytes, spleen, kidney, testis and muscle (obtained from Origene). The DD11 molecular marker was further tested using cDNA populations derived from a more comprehensive panel of 22 human tissue types. These are as follows:

Adrenal gland	pooled from 62 donors
Bone marrow	pooled from 7 donors
Brain, cerebellum	pooled from 24 donors
Brain, whole	pooled from one donor
Colon*	pooled from one donor
Foetal brain	pooled from 59 donors
Foetal liver	pooled from 63 donors
Heart	pooled from one donor
Kidney	pooled from one donor
Liver	pooled from one donor
Lung	pooled from one donor
Placenta	pooled from 7 donors
Prostate	pooled from 47 donors
Salivary gland	pooled from 24 donors
Skeletal muscle	pooled from 2 donors
Small intestine*	pooled from one donor
Spleen	pooled from 14 donors
Testis	pooled from 19 donors
Thymus	pooled from 9 donors
Thyroid gland	pooled from 65 donors
Trachea	pooled from ? donors
Uterus	pooled from 10 donors

Note that the majority of these samples were part of the Human Total RNA panel II (Clontech), but two samples, marked with asterisks, were obtained as tissue chunks from Pathlore (Peterborough Hospital Tissue Bank) and processed at Randox Laboratories Ltd.

In addition, assays were performed on a range of ethically approved human tumour samples, as obtained through Pathlore. cDNA representative of tumours from ovary, testis, stomach, liver, lung, bladder, colon and pancreas were tested against both β-actin and DD11 by real-time PCR.

In conjunction with novel marker expression analysis, each matched pair of breast tissues was subjected to molecular signature analysis. This entailed using a suite of primers specific to a number of pre-published breast cancer

molecular markers in semi-quantitative RT-PCR against each tissue cDNA. The relationship between each molecular marker was determined and tabulated for each sample and used as a reference, against which the novel markers could be compared. This was with the aim of sub-classifying the tumour types to enable the association of novel markers against such sub-types, increasing the power of the diagnostic marker considerably.

Results and Discussion.

Using differential display, a gene fragment, termed DD11, derived from cDNA populations of matched tissue from a breast cancer donor, was observed to have significant up-regulation in the tumour cDNA population in comparison to the corresponding normal tissue cDNA. This 171-nucleotide product (Figure 1) was confirmed as differentially expressed by reverse Southern dot blots. Sequence analysis followed by database interrogation determined that DD11 was not homologous to known genes or proteins in the EMBL and SWISSPROT databases, respectively, so was regarded as potentially novel. It was, however, 100% homologous, after removal of the poly-A tail, to a clone from chromosome 8 of the human genome (Figure 2).

This fragment was further screened using cDNA populations derived from a number of matched breast tumour tissues donated by other patients. Of the donor samples screened, 6 out of 9 exhibited notable increases in expression, confirming DD11 to be a putative molecular marker for the presence of breast tumour (Figure 3). This analysis was substantiated by the molecular signature analysis of all currently available matched breast tissue samples, as follows;

Increased in tumour	10	52.6%
Increased in normal	1	5.3%
No discernable difference	7	36.8%
No expression evident	1	5.3%
Totals	19	100%

To facilitate further analysis, 5'-RACE was employed to extend the fragment to include the full open reading frame (ORF) of the gene, plus any 5' non-coding sequence. Using this technique, a presumed full-length product of 513 nucleotides was derived, which on subsequent database interrogation,

confirmed the previous homology to human chromosome 8, being 100% homologous over the full length of the sequence (513/513) (Figure 4). From this sequence, all 6 amino acid reading frames were generated and a putative, small ORF was found in the +2 frame, comprising 48 amino acids, including the stop codon (Figure 5). This small protein failed to reveal a high homology to any known proteins in the SWALL database, so is assumed to be novel.

To determine organ specificity, cDNA populations from 8 non-breast human tissues were tested against the DD11 primers, in addition to a matched pair of cDNAs from a breast cancer donor. The same samples were also tested using primers from the constitutive housekeeping gene, β-actin, as a positive control and to calibrate the templates for semi-quantitative PCR analysis. The β-actin product was strongly amplified in all cDNA populations studied, whereas the DD11 product was only detected in the breast tumour sample (Figure 6). This provided further evidence that this novel gene could be a very powerful molecular marker for the presence of a breast tumour.

This molecular marker was further tested using cDNA populations derived from a panel of 22 human tissue types, both by conventional and real-time PCR analysis. Of those tested, using the Opticon II real-time thermal cycler (MJ Research), DD11 was only detected in samples from placenta and testis (data not shown). In addition, assays were performed on a range of ethically approved human tumour samples, as obtained through Pathlore, to ascertain whether the marker was breast tumour specific or a less specific tumour marker. cDNA representative of tumours from ovary, testis, colon, stomach, liver, lung, bladder and pancreas were tested against both β-actin and DD11. Of these, DD11 was only detected at a significant level in cDNA derived from testis tumour (data not shown). The products from these PCR amplifications were verified as DD11 by direct sequencing. Initial screening of a testis sample from another source (Origene), however, failed to detect this product (see figure 6).

Conventional PCR amplification on a standard thermal cycler, using the combined panels of 22 normal human tissue cDNAs and the 8 tumour cDNA populations, confirmed DD11 to be specific to a very limited number of tissue types (Figure 7). As in the real-time PCR analysis, the testis cDNA population

expressed DD11, as did the testis tumour. The only other population showing significant expression of this marker was from the uterus sample. The placenta sample showed a low level of expression, but this was not to the same scale as the testis and uterus samples. Low levels of product were also found in some other tissue samples, but these were considered negligible. This would indicate that of all the samples tested, DD11 is only strongly expressed in those tissues under the influence of reproductive hormones. Specifically, the tissues derived from breast and associated tumour, testis and associated tumour and uterus. Placenta tissue also expresses this marker to a lesser extent.

It should also be noted that this molecular marker is not expressed in all breast and tumour samples tested, and so it may be useful for sub-classification of the breast tumour type. Comparison of the expression profiles of DD11 in the tissue samples, against the molecular signatures may reveal associations between this marker and other pre-published breast cancer markers, which have been linked to disease classification and prognosis.

For reference, it is important to point out that DD11 compares very favourably with some of the most highly regarded "standard" breast cancer markers, such as Oestrogen receptor (ERα) and human epidermal growth factor receptor (c-ErbB-2). This is evident both in the molecular signature analysis of all matched breast cancer tissue samples, where expression is similar in both samples from the same patient in many cases and using the target-specific primers against our in-house panel of 30 cDNA populations from human normal and tumour tissue. Two examples of these, namely ERα and c-ErbB-2, are shown in Figures 8 and 9, respectively. In addition, the screening of c-ErbB-2 against a selection of our matched samples is given in Figure 10. In all cases, amplified targets have been verified by sequence analysis.

References.

DeRisi, J. L., Iyer, V. R. and Brown, P. O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. Science. 278: 680-686.

Liang, P. and Pardee, A. B. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase reaction. Science. 257: 967-971.

Pritzker, K. P. 2002 Cancer biomarkers: easier said than done. Clin. Chem. 2002 Aug; 48(8):1147-50.

Salodof MacNeil. 2001. From genes to proteins: The FLEXgene consortium. HMS Beagle. 112: on-line journal.

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. Sep 11;98(19):10869-74.

Srinivas PR, Verma M, Zhao Y, Srivastava S. 2002. Proteomics for cancer biomarker discovery. Clin Chem. Aug;48(8):1160-9.

Sturtevant, J. Applications of differential-display reverse transcription-PCR to molecular pathogenesis and medical mycology. Clin Microbiol Rev. 2000 Jul;13(3):408-27.

Yousef et al., 2002 Yousef GM, Scorilas A, Kyriakopoulou LG, Rendl L, Diamandis M, Ponzone R, Biglia N, Giai M, Roagna R, Sismondi P, Diamandis EP. Human kallikrein gene 5 (KLK5) expression by quantitative PCR: an independent indicator of poor prognosis in breast cancer. Clin Chem. 2002 Aug;48(8):1241-50.

Zong, Q., Schummer, M., Hood, L. and Morris, D. R. 1999. Messenger RNA translation state: the second dimension of high-throughput expression screening. Proc. Natl. Acad. Sci. 96: 10632-10636.

CLAIMS

1. A method for the detection of the presence of or the risk of cancer in a patient comprising the steps of:

(i) isolating a sample of the patient's genome; and

(ii) detecting the presence or expression of the gene characterised by the nucleotide sequence identified as SEQ ID No. 1,

wherein the presence or expression of the gene indicates the presence of or the risk of cancer.

2. A method according to claim 1, wherein the genome sample is obtained from breast tissue, the uterus or testis.

3. A method according to claim 1 or claim 2, wherein the cancer is breast cancer.

4. A method according to any preceding claim, wherein detection is carried out by amplifying the gene using the polymerase enzyme.

5. An isolated polynucleotide comprising the nucleotide sequence identified herein as SEQ ID No. 1, or its complement, or a polynucleotide of at least 15 consecutive nucleotides that hybridises to the sequence (or its complement) under stringent hybridising conditions.

6. Use of a polynucleotide according to claim 5, in an *in vitro* diagnostic assay to test for the risk of cancer in a patient.

7. Use according to claim 6, wherein the cancer is breast cancer.

8. A peptide comprising the sequence identified herein as SEQ ID No. 2, or a fragment thereof of at least 10 consecutive amino acid residues.

9. An antibody having an affinity of at least 10^{-6}M for the peptide of claim 8.

10. Use of a polynucleotide that hybridises with or inhibits the expression of an endogenous gene that comprises the polynucleotide according to claim 5, in the manufacture of a medicament for the treatment of cancer, in particular breast cancer.

Figure 1. Nucleotide sequence of DD11, including the poly-A tail.

```
GATTAGTCCTCAGGCATCTCCCTGCTCTGAGCTGAGGGGTGTGGTAGTGTTGAAAGTTGCAGTGCTCTGATCACGTGGTTGGTTCG
ACTGGTAACTGGTCCCTCTCTGGCAAGAGCCACCTCATCAGTATCAACTCAGGAATGCTGGAAATCATTTTATGAAAAAAAAAAA
```

Figure 2. Results of database searches on DD11.

```
Query=(160 letters)

Database:  embl

                                                                    Smallest
                                                                      Sum
                                                              High  Probability
Sequences producing High-scoring Segment Pairs:               Score  P(N)       N

EM_HUM:AC107959 AC107959.8 Homo sapiens chromosome 8, clo...   800  3.5e-28    1
EM_HUM:AC100861 AC100861.2 Homo sapiens chromosome 8, clo...   321  1.5e-06    1
EM_HUM:AF165424 AF165424.5 Homo sapiens chromosome 8 clon...   305  7.9e-06    1
EM_HUM:AC103760 AC103760.2 Homo sapiens chromosome 8, clo...   305  7.9e-06    1

>EM_HUM:AC107959 AC107959.8 Homo sapiens chromosome 8, clone RP11-875O11,
         complete sequence.
       Length = 211,291

 Plus Strand HSPs:

 Score = 800 (126.1 bits), Expect = 3.5e-28, P = 3.5e-28
 Identities = 160/160 (100%), Positives = 160/160 (100%), Strand = Plus / Plus

Query:      1 GATTAGTCCTCAGGCATCTCCCTGCTCTGAGCTGAGGGGTGTGGTAGTGTTGAAAGTTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 157022 GATTAGTCCTCAGGCATCTCCCTGCTCTGAGCTGAGGGGTGTGGTAGTGTTGAAAGTTGC 157081

Query:     61 AGTGCTCTGATCACGTGGTTGGTTCGACTGGTAACTGGTCCCTCTCTGGCAAGAGCCACC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 157082 AGTGCTCTGATCACGTGGTTGGTTCGACTGGTAACTGGTCCCTCTCTGGCAAGAGCCACC 157141

Query:    121 TCATCAGTATCAACTCAGGAATGCTGGAAATCATTTTATG 160
              |||||||||||||||||||||||||||||||||||||||
Sbjct: 157142 TCATCAGTATCAACTCAGGAATGCTGGAAATCATTTTATG 157181
```

Figure 3. Tumour specificity of DD11, as revealed by screening a number of matched breast tumour cDNA populations from breast cancer donors. The housekeeping gene, B-actin was used to standardise the cDNA populations. T represents tumour tissue cDNA whereas M represents co-excised mammary tissue cDNA from the same donor.

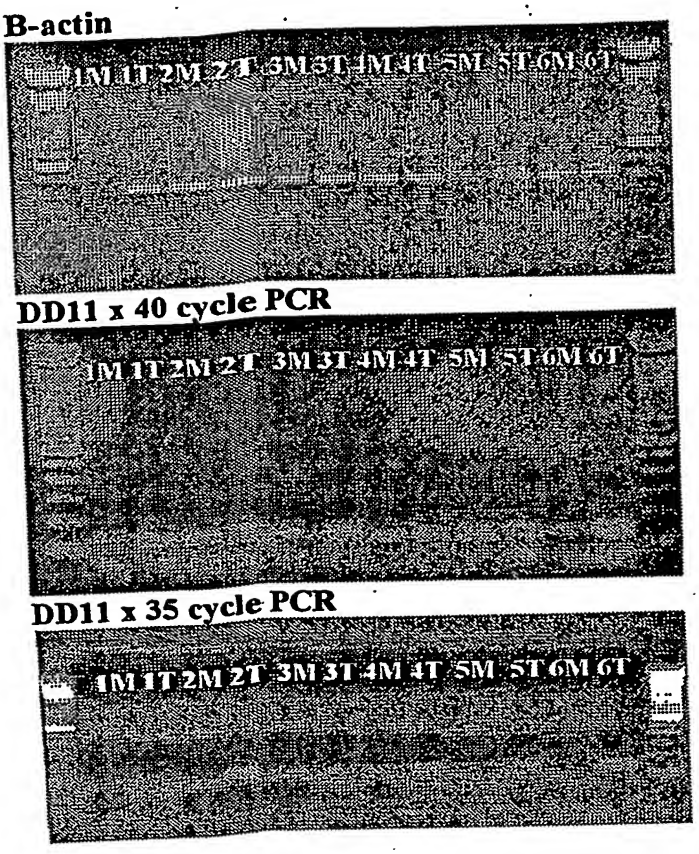


Figure 4. Presumed full-length sequence of DDI1 and results of the subsequent database interrogation.

```
   1  TTTATGGTCA TAAGCTTAGA AAATCCTTTG CCCAACATAA AATAAGAGAA CTCTAATTTC   60
  61  TTAGGGAGAT TTTTATTAAA TGATTAGATT TGTAGCATAT AGTTGTATAA AATAAGATGA  120
 121  ACTCTAATTT CTTAGGGAGG TTTTATTAAA TGATTAGATT TGTAGCATAT CATCGTGTAA  180
 181  AGTACATGGA CATTATTTTT GATATAGAAA GTGTAGTGTT CCCCTTCATT GTTCTGAGTT  240
 241  ACTCTCATCT GTCCAACCCC AGCGAGCCAC TGATTATTCC CTTTCTCTGA ACTTTGTGGT  300
 301  GTTTATGGAA GCTTCATTCC GTAGCACGAA GGCGTCAATC ATTAATCTCG GGTGATTAGT  360
 361  CCTCAGGCAT CTCCCTGCTC TGAGCTGAGG GGTGTGGTAG TGTTGAAAGT TGCAGTGCTC  420
 421  TGATCACGTG GTTGGTTCGA CTGGTAACTG GTCCCTCTCT GGCAAGAGCC ACCTCATCAG  480
 481  TATCAACTCA GGAATGCTGG AAATCATTTT ATG                               513

Query=  (513 letters)
Database:  embl
           2,705,345 sequences; 8,456,263,008 total letters.
Searching....10....20....30....40....50....60....70....80....90....100% done
```

Sequences producing High-scoring Segment Pairs:		High Score	Smallest Sum Probability P(N)	N
EM_HUM:AC107959	AC107959.8 Homo sapiens chromosome 8, clo...	2565	7.0e-108	1
EM_HUM:AC100861	AC100861.2 Homo sapiens chromosome 8, clo...	1368	8.2e-54	1
EM_HUM:AP003071	AP003071.3 Homo sapiens genomic DNA, chro...	409	1.7e-10	1

```
>EM_HUM:AC107959 AC107959.8 Homo sapiens chromosome 8, clone RP11-875O11,
          complete sequence.
        Length = 211,291

 Score = 2565 (390.9 bits), Expect = 7.0e-108, P = 7.0e-108
 Identities = 513/513 (100%), Positives = 513/513 (100%), Strand = Plus / Plus

Query:      1 TTTATGGTCATAAGCTTAGAAAATCCTTTGCCCAACATAAAATAAGAGAACTCTAATTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 156669 TTTATGGTCATAAGCTTAGAAAATCCTTTGCCCAACATAAAATAAGAGAACTCTAATTTC 156728

Query:     61 TTAGGGAGATTTTTATTAAATGATTAGATTTGTAGCATATAGTTGTATAAAATAAGATGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 156729 TTAGGGAGATTTTTATTAAATGATTAGATTTGTAGCATATAGTTGTATAAAATAAGATGA 156788

Query:    121 ACTCTAATTTCTTAGGGAGGTTTTATTAAATGATTAGATTTGTAGCATATCATCGTGTAA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 156789 ACTCTAATTTCTTAGGGAGGTTTTATTAAATGATTAGATTTGTAGCATATCATCGTGTAA 156848

Query:    181 AGTACATGGACATTATTTTTGATATAGAAAGTGTAGTGTTCCCCTTCATTGTTCTGAGTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 156849 AGTACATGGACATTATTTTTGATATAGAAAGTGTAGTGTTCCCCTTCATTGTTCTGAGTT 156908

Query:    241 ACTCTCATCTGTCCAACCCCAGCGAGCCACTGATTATTCCCTTTCTCTGAACTTTGTGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 156909 ACTCTCATCTGTCCAACCCCAGCGAGCCACTGATTATTCCCTTTCTCTGAACTTTGTGGT 156968

Query:    301 GTTTATGGAAGCTTCATTCCGTAGCACGAAGGCGTCAATCATTAATCTCGGGTGATTAGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 156969 GTTTATGGAAGCTTCATTCCGTAGCACGAAGGCGTCAATCATTAATCTCGGGTGATTAGT 157028

Query:    361 CCTCAGGCATCTCCCTGCTCTGAGCTGAGGGGTGTGGTAGTGTTGAAAGTTGCAGTGCTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 157029 CCTCAGGCATCTCCCTGCTCTGAGCTGAGGGGTGTGGTAGTGTTGAAAGTTGCAGTGCTC 157088

Query:    421 TGATCACGTGGTTGGTTCGACTGGTAACTGGTCCCTCTCTGGCAAGAGCCACCTCATCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 157089 TGATCACGTGGTTGGTTCGACTGGTAACTGGTCCCTCTCTGGCAAGAGCCACCTCATCAG 157148

Query:    481 TATCAACTCAGGAATGCTGGAAATCATTTTATG 513
              ||||||||||||||||||||||||||||||||
Sbjct: 157149 TATCAACTCAGGAATGCTGGAAATCATTTTATG 157181
```

Figure 5. Plus 3 amino acid reading frame for DD11 showing the putative small ORF indicated in bold text.

```
  3     YGHKLRKSFA QHKIREL*FL REIFIK*LDL *HIVV*NKMN SNFLGRFY*M      152
153     IRFVAYHRVK YMDIIFDIES VVFPFIVLSY SHLSNPSEPL IIPFL*TLWC      302
303     LWKLHSVARR RQSLISGD*S SGISLL*AEG CGSVESCSAL ITWLVRLVTG      452
453     PSLARATSSV STQECWKSFY                                       512
```

Figure 6. Expression analysis of β-actin and DD11 on a number of cDNA populations derived from non-breast tissues and cDNA populations from a matched pair of breast tumour and normal tissues from a breast cancer donor. This shows expression of DD11 only in the breast tumour tissue, whereas β-actin is expressed strongly in all tissue samples tested.

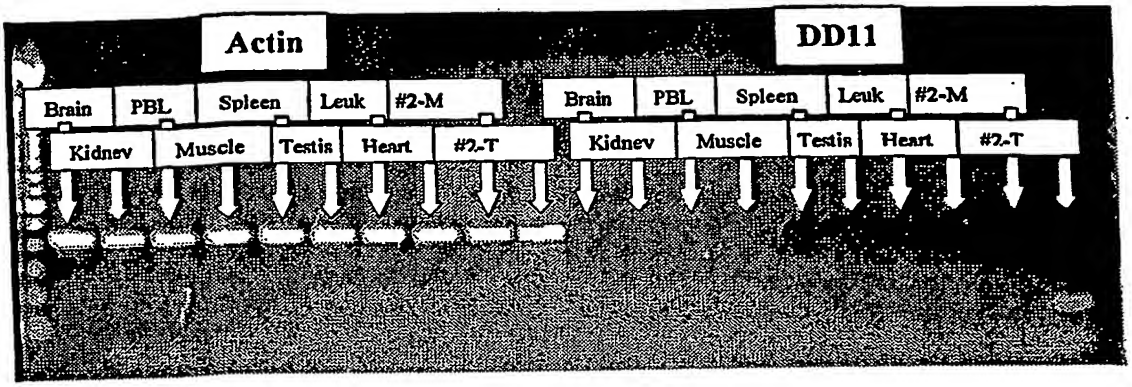
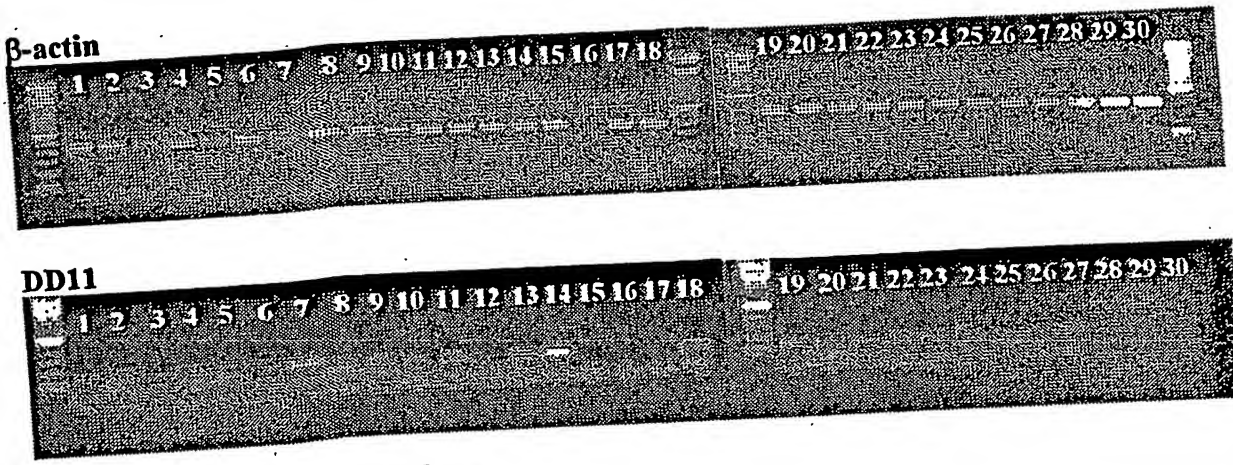


Figure 7. β-actin and DD11 checked against the panels of cDNA derived from human normal and tumour tissue samples. Strong expression of DD11 is limited to testis tumour, testis and, to a lesser extent, uterus tissue samples.

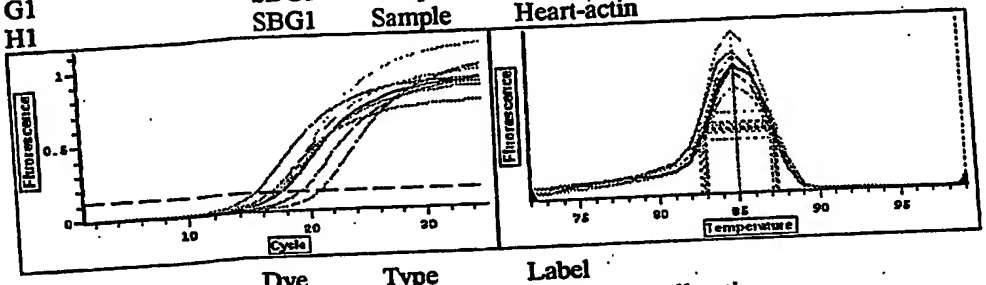


Order of cDNA templates

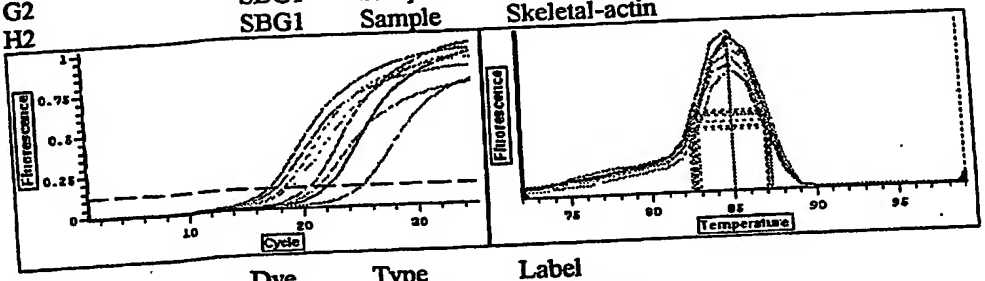
1. Pancreas Tumour
2. Lung Tumour
3. Liver Tumour
4. Ovary Tumour
5. Stomach Tumour
6. Bladder Tumour
7. √ Testis Tumour
8. Colon Tumour
9. Prostate
10. Brain-Cerebellum
11. Kidney
12. Heart
13. Small Intestine
14. √ Testis
15. Skeletal muscle
16. Colon
17. Thyroid
18. √ Uterus
19. Placenta
20. Trachea
21. Foetal brain
22. Bone marrow
23. Adrenal gland
24. Thymus
25. Lung
26. Brain-Whole
27. Foetal liver
28. Spleen
29. Liver
30. Salivary gland

Figure 8. Screening of the standard breast cancer molecular marker, ERα, against the panel of 22 human tissue and tumour cDNAs, shown as real-time data

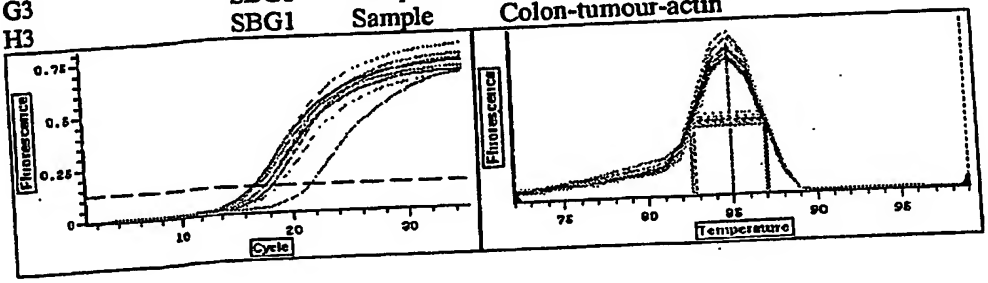
Well	Dye	Type	Label	C(T)
A1	SBG1	Sample	Adrenal-actin	19.148
B1	SBG1	Sample	Bone marrow-actin	16.557
C1	SBG1	Sample	Brain-Cereb-actin	17.542
D1	SBG1	Sample	Brain-Whole-actin	16.769
E1	SBG1	Sample	Colon-actin	15.263
F1	SBG1	Sample	Foetal Brain-actin	17.315
G1	SBG1	Sample	Foetal Liver-actin	17.149
H1	SBG1	Sample	Heart-actin	20.783



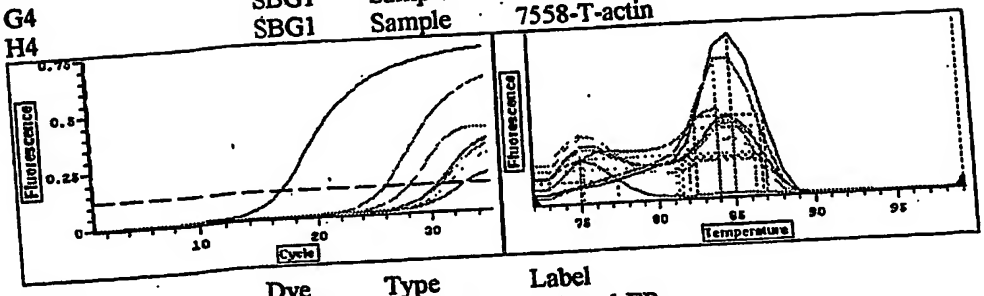
Well	Dye	Type	Label	C(T)
A2	SBG1	Sample	Intestine-small-actin	17.041
B2	SBG1	Sample	Kidney-actin	21.248
C2	SBG1	Sample	Liver-actin	22.125
D2	SBG1	Sample	Lung-actin	18.816
E2	SBG1	Sample	Placenta-actin	25.145
F2	SBG1	Sample	Prostate-actin	17.603
G2	SBG1	Sample	Salivary-actin	18.318
H2	SBG1	Sample	Skeletal-actin	20.463



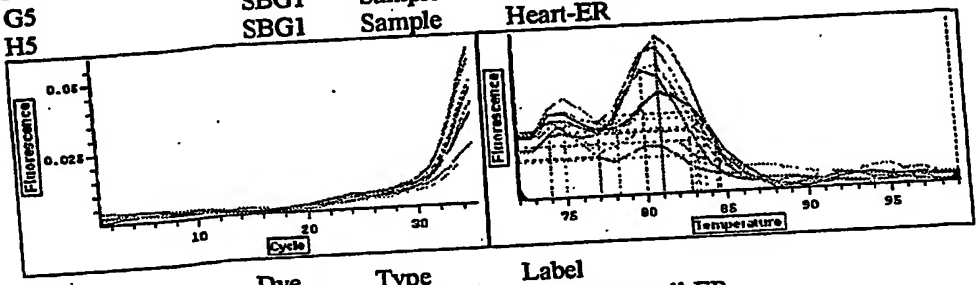
Well	Dye	Type	Label	C(T)
A3	SBG1	Sample	Spleen-actin	17.512
B3	SBG1	Sample	Testis-actin	17.514
C3	SBG1	Sample	Thymus-actin	16.559
D3	SBG1	Sample	Thyroid-actin	16.799
E3	SBG1	Sample	Trachea-actin	18.124
F3	SBG1	Sample	Uterus-actin	16.971
G3	SBG1	Sample	Bladder-tumour-actin	18.612
H3	SBG1	Sample	Colon-tumour-actin	21.413



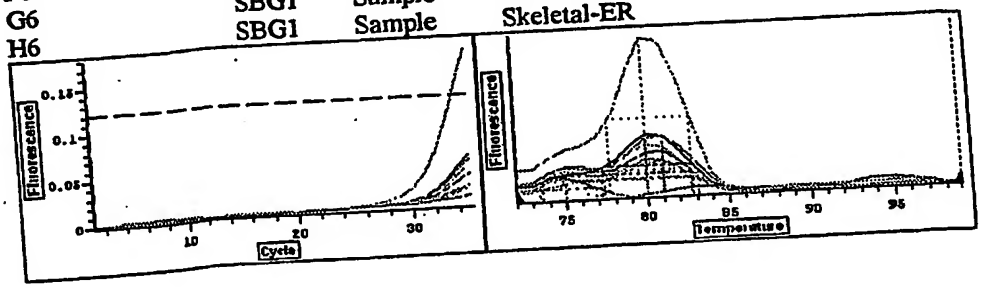
Well	Dye	Type	Label	C(T)
A4	SBG1	Sample	Liver-tumour-actin	32.978
B4	SBG1	Sample	Lung-tumour-actin	30.809
C4	SBG1	Sample	Ovary-tumour-actin	16.856
D4	SBG1	Sample	Pancreas-tumour-actin	30.537
E4	SBG1	Sample	Stomach-tumour-actin	30.59
F4	SBG1	Sample	Testis-tumour-actin	28.122
G4	SBG1	Sample	7558-M-actin	31.613
H4	SBG1	Sample	7558-T-actin	26.03



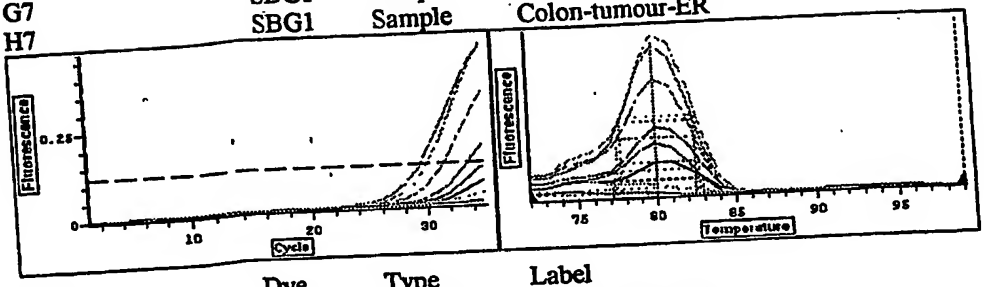
Well	Dye	Type	Label	C(T)
A5	SBG1	Sample	Adrenal-ER	None
B5	SBG1	Sample	Bone marrow-ER	None
C5	SBG1	Sample	Brain-Cereb-ER	None
D5	SBG1	Sample	Brain-Whole-ER	None
E5	SBG1	Sample	Colon-ER	None
F5	SBG1	Sample	Foetal Brain-ER	None
G5	SBG1	Sample	Foetal Liver-ER	None
H5	SBG1	Sample	Heart-ER	None



Well	Dye	Type	Label	C(T)
A6	SBG1	Sample	Intestine-small-ER	None
B6	SBG1	Sample	Kidney-ER	None
C6	SBG1	Sample	Liver-ER	None
D6	SBG1	Sample	Lung-ER	None
E6	SBG1	Sample	Placenta-ER	33.608
F6	SBG1	Sample	Prostate-ER	None
G6	SBG1	Sample	Salivary-ER	None
H6	SBG1	Sample	Skeletal-ER	None



Well	Dye	Type	Label	C(T)
A7	SBG1	Sample	Spleen-ER	None
B7	SBG1	Sample	Testis-ER	31.447
C7	SBG1	Sample	Thymus-ER	None
D7	SBG1	Sample	Thyroid-ER	30.03
E7	SBG1	Sample	Trachea-ER	33.52
F7	SBG1	Sample	Uterus-ER	29.325
G7	SBG1	Sample	Bladder-tumour-ER	None
H7	SBG1	Sample	Colon-tumour-ER	None



Well	Dye	Type	Label	C(T)
A8	SBG1	Sample	Liver-tumour-ER	None
B8	SBG1	Sample	Lung-tumour-ER	None
C8	SBG1	Sample	Ovary-tumour-ER	29.145
D8	SBG1	Sample	Pancreas-tumour-ER	None
E8	SBG1	Sample	Stomach-tumour-ER	None
F8	SBG1	Sample	Testis-tumour-ER	None
G8	SBG1	Sample	7558-M-ER	None
H8	SBG1	Sample	7558-T-ER	31.058

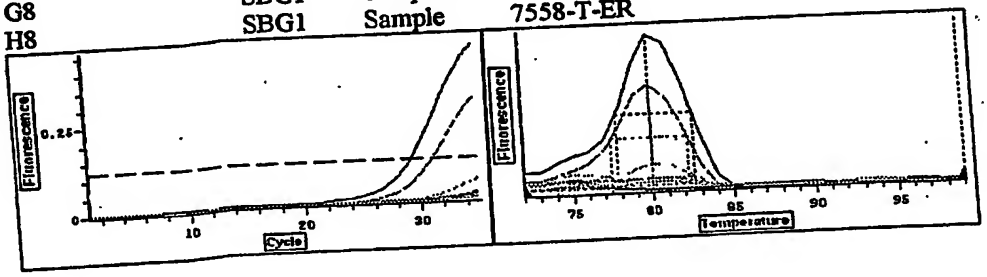
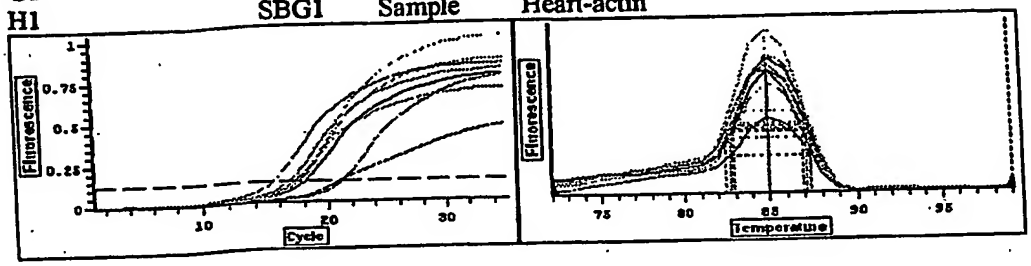
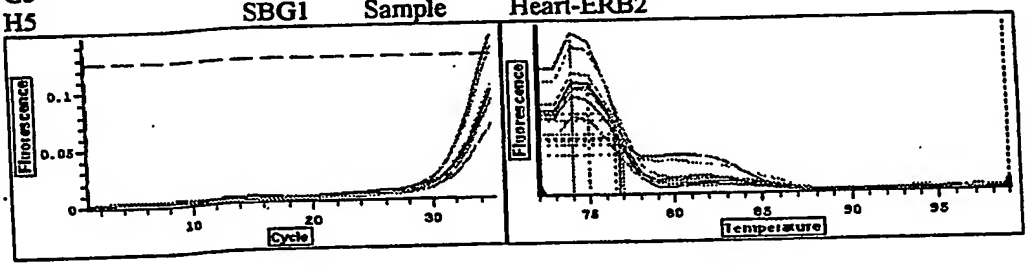


Figure 9. Screening of the standard breast cancer molecular marker, c-ErbB-2, against the panel of 22 human tissue and tumour cDNAs, shown as real-time data.

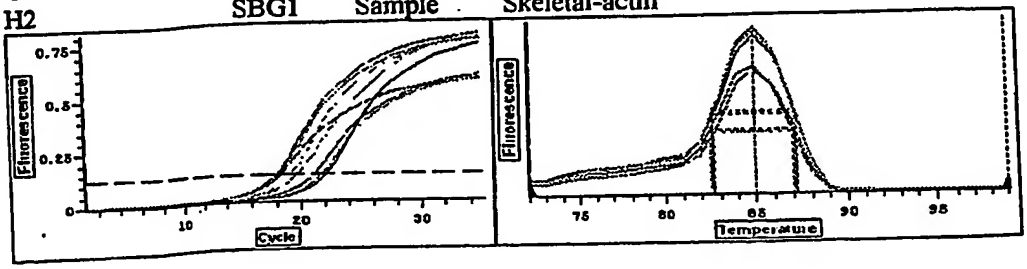
Well	Dye	Type	Label	C(T)
A1	SBG1	Sample	Adrenal-actin	21.143
B1	SBG1	Sample	Bone marrow-actin	17.294
C1	SBG1	Sample	Brain-Cereb-actin	18.104
D1	SBG1	Sample	Brain-Whole-actin	17.534
E1	SBG1	Sample	Colon-actin	15.433
F1	SBG1	Sample	Foetal Brain-actin	16.772
G1	SBG1	Sample	Foetal Liver-actin	17.541
H1	SBG1	Sample	Heart-actin	21.257



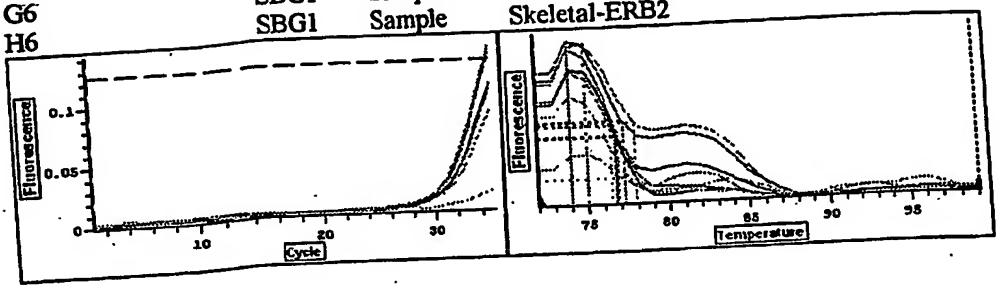
Well	Dye	Type	Label	C(T)
A5	SBG1	Sample	Adrenal-ERB2	34.416
B5	SBG1	Sample	Bone marrow-ERB2	34.745
C5	SBG1	Sample	Brain-Cereb-ERB2	None
D5	SBG1	Sample	Brain-Whole-ERB2	None
E5	SBG1	Sample	Colon-ERB2	None
F5	SBG1	Sample	Foetal Brain-ERB2	None
G5	SBG1	Sample	Foetal Liver-ERB2	None
H5	SBG1	Sample	Heart-ERB2	None



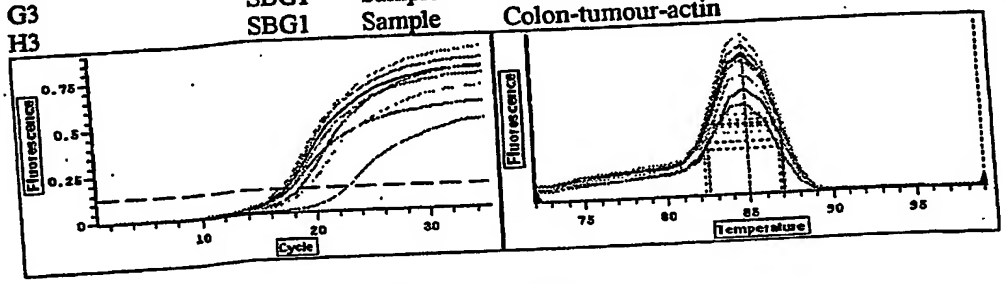
Well	Dye	Type	Label	C(T)
A2	SBG1	Sample	Intestine-small-actin	18.352
B2	SBG1	Sample	Kidney-actin	21.784
C2	SBG1	Sample	Liver-actin	22.140
D2	SBG1	Sample	Lung-actin	19.684
E2	SBG1	Sample	Placenta-actin	17.995
F2	SBG1	Sample	Prostate-actin	18.189
G2	SBG1	Sample	Salivary-actin	19.161
H2	SBG1	Sample	Skeletal-actin	21.121



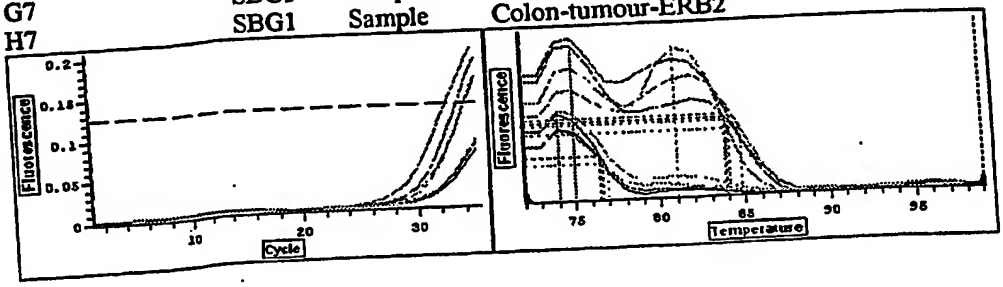
Well	Dye	Type	Label	C(T)
A6	SBG1	Sample	Intestine-small-ERB2	34.782
B6	SBG1	Sample	Kidney-ERB2	None
C6	SBG1	Sample	Liver-ERB2	None
D6	SBG1	Sample	Lung-ERB2	None
E6	SBG1	Sample	Placenta-ERB2	34.614
F6	SBG1	Sample	Prostate-ERB2	None
G6	SBG1	Sample	Salivary-ERB2	34.794
H6	SBG1	Sample	Skeletal-ERB2	



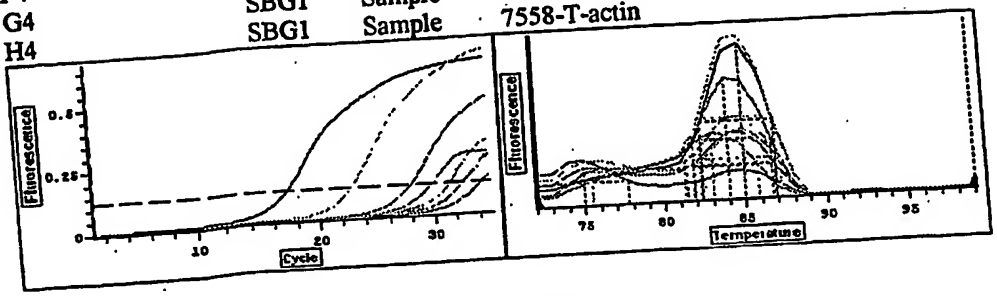
Well	Dye	Type	Label	C(T)
A3	SBG1	Sample	Spleen-actin	17.643
B3	SBG1	Sample	Testis-actin	17.527
C3	SBG1	Sample	Thymus-actin	17.235
D3	SBG1	Sample	Thyroid-actin	16.800
E3	SBG1	Sample	Trachea-actin	18.473
F3	SBG1	Sample	Uterus-actin	17.049
G3	SBG1	Sample	Bladder-tumour-actin	19.166
H3	SBG1	Sample	Colon-tumour-actin	22.747



Well	Dye	Type	Label	C(T)
A7	SBG1	Sample	Spleen-ERB2	None
B7	SBG1	Sample	Testis-ERB2	34.041
C7	SBG1	Sample	Thymus-ERB2	None
D7	SBG1	Sample	Thyroid-ERB2	None
E7	SBG1	Sample	Trachea-ERB2	33.373
F7	SBG1	Sample	Uterus-ERB2	33.958
G7	SBG1	Sample	Bladder-tumour-ERB2	None
H7	SBG1	Sample	Colon-tumour-ERB2	32.476



Well	Dye	Type	Label	C(T)
A4	SBG1	Sample	Liver-tumour-actin	34.271
B4	SBG1	Sample	Lung-tumour-actin	31.338
C4	SBG1	Sample	Ovary-tumour-actin	17.152
D4	SBG1	Sample	Pancreas-tumour-actin	22.753
E4	SBG1	Sample	Stomach-tumour-actin	32.659
F4	SBG1	Sample	Testis-tumour-actin	29.761
G4	SBG1	Sample	7558-M-actin	32.761
H4	SBG1	Sample	7558-T-actin	27.948



Well	Dye	Type	Label	C(T)
A8	SBG1	Sample	Liver-tumour-ERB2	34.812
B8	SBG1	Sample	Lung-tumour-ERB2	32.515
C8	SBG1	Sample	Ovary-tumour-ERB2	30.024
D8	SBG1	Sample	Pancreas-tumour-ERB2	None
E8	SBG1	Sample	Stomach-tumour-ERB2	None
F8	SBG1	Sample	Testis-tumour-ERB2	34.800
G8	SBG1	Sample	7558-M-ERB2	None
H8	SBG1	Sample	7558-T-ERB2	None

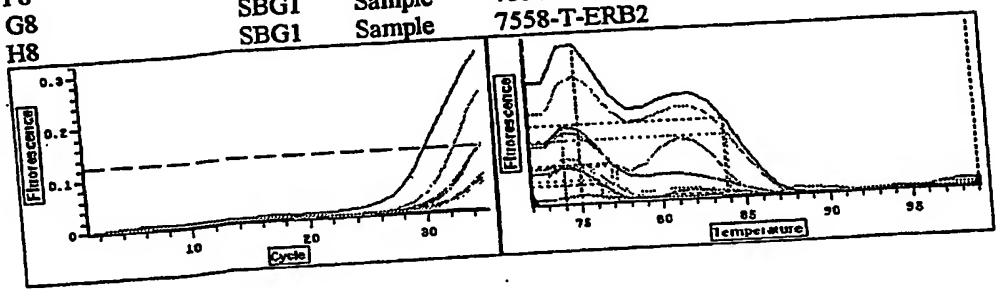
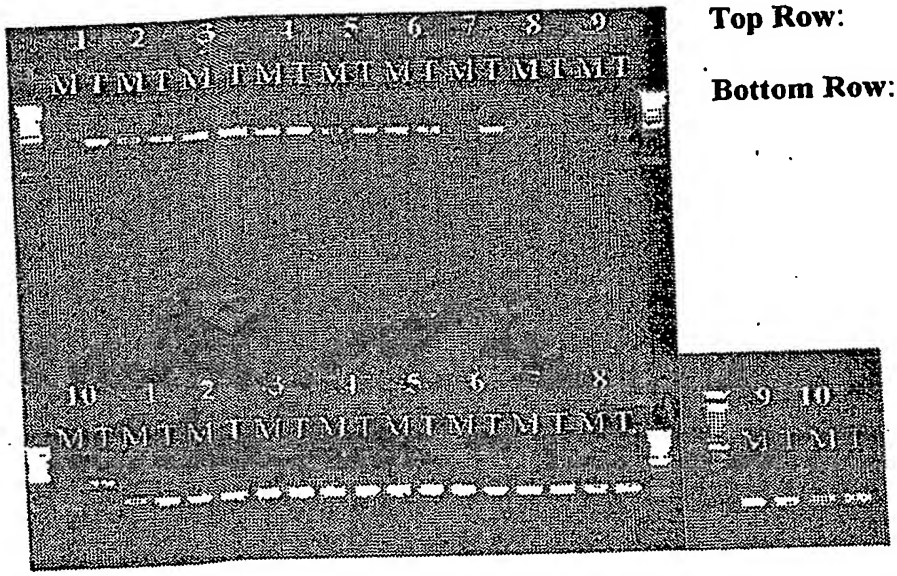


Figure 10. Screening of the standard breast cancer molecular marker, c-ErbB-2, against 10 of the matched normal and tumour breast tissue cDNAs. β-actin expression is also shown, for confirmation of template calibration and integrity. Note that increased tumour expression of this gene is only evident in 4 of the 10 matched tissues.



Top Row: c-ErbB-2

Bottom Row: β-actin